

A NEW DIRECT PROOF OF THE CENTRAL LIMIT THEOREM

VLADIMIR DOBRIĆ AND PATRICIA GARMIRIAN

ABSTRACT. We prove the Central Limit Theorem (CLT) from the definition of weak convergence using the Haar wavelet basis, calculus, and elementary probability. The use of the Haar basis pinpoints the role of $L^2([0, 1])$ in the CLT as well as the assumption of finite variance. We estimate the rate of convergence and prove strong convergence off the tails.

1. INTRODUCTION

The Central Limit Theorem (CLT) is one of the most fundamental theorems of probability theory. The theorem states that standardized sums of i.i.d. random variables having finite variance converge weakly to the standard normal distribution. As early as the 1770s, mathematicians were searching for the “central limit,” trying to establish the correct conditions for convergence in distribution and the formula for the limiting distribution. The connection between convergence in distribution and characteristic functions was established in the 1920s by Lévy.

The connection between convergence in distribution and weak convergence was established in the late 1940s (see [6]). For a measurable space $(S, \mathcal{B}(S))$, where S is a Polish space, a sequence of measures μ_n converges weakly to μ provided that for each bounded, continuous function $f : S \rightarrow \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \int_S f(x) d\mu_n(x) = \int_S f(x) d\mu(x).$$

The advantage of using this definition is that it may serve as a stepping-stone to extending the CLT to random variables having values in more general spaces.

A stronger type of convergence for measures than “weak” convergence is “strong” convergence. For a measurable space $(S, \mathcal{B}(S))$, a sequence of measures μ_n converges strongly to μ provided that for each set $A \in \mathcal{B}(S)$,

$$\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A).$$

We show that the type of convergence in the CLT is in fact strong off the tails.

In 1935, both Feller, [5], and Lévy (independently), [8], [9], proved the Central Limit Theorem (CLT) using characteristic functions. As the CLT is a fundamental theorem in probability theory, since the time of Donsker many experts have believed that there should be a direct proof (see [6] for example). Also, there is obvious interest in determining rates and constants of convergence and these characteristic function proofs gave no information about these issues.

Since 1935, there have been a number of more elementary or direct proofs of the CLT which do not use characteristic functions, e.g., [1], [2], [3], [4], [10] and [11]. The last two prove the CLT directly from the definition of weak convergence and [2] and (taken together) [1], [3], [4], do give a rate of convergence of $n^{-1/2}$ as well as a constant of convergence. However, all of these proofs involve a hypothesis stronger than the optimal hypothesis of Feller and Lévy, namely finite variance. The hypothesis of [1], [2], [3], [4] and [11] is finite third moment, while the hypothesis of [10] is continuous second derivative of the function f in the definition of weak convergence.

Date: 24 July 2015.

2010 Mathematics Subject Classification. Primary 60F05 ; Secondary 42C40, 28A33.

Key words and phrases. Central Limit Theorem, Haar basis.

Garmirian dedicates this paper to the memory of her dearly departed mentor, Vladimir Dobrić, and acknowledges the helpful comments of Lee Stanley, Rob Neel, and Daniel Conus.

We present a proof, directly from the definition of weak convergence, avoiding characteristic functions, which is optimal in terms of hypothesis (as in [5], [8], [9]). We elaborate on the next complex of notions in what follows, but for now, it will suffice to say that the proof involves multinomial approximations to the initial sum of random variables, and identifies “tails” of both these multinomials and the Gaussian. We obtain strong convergence off these tails, i.e. the sum of the absolute values of the differences between the multinomial and Gaussian probabilities converges to zero. Our proof also provides estimates of the rate and constant of convergence off these tails; the former is comparable to the rates of [2] and ([1], [3], [4]).

The translation into the language of random variables of the above definition of weak convergence is established by letting $\mu_n = P \circ X_n^{-1}$. A sequence of random variables (X_n) converges weakly to a random variable X if for each bounded, continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$\lim_{n \rightarrow \infty} E(f(X_n)) = E(f(X)),$$

and it is this translation that we use in what follows. A sequence of random variables X_n on a measurable space $(S, \mathcal{B}(S))$ converges strongly to a random variable X if for each $A \in \mathcal{B}(S)$,

$$\lim_{n \rightarrow \infty} P(X_n \in A) = P(X \in A).$$

In the case where X_n and X are discrete with range set J , by the triangle inequality, it is sufficient to show that

$$\lim_{n \rightarrow \infty} \sum_{j \in J} |P(X_n = j) - P(X = j)| = 0.$$

In fact, approximating the initial sum of random variables and the Gaussian by discrete versions, the preceding statement holds for the sequences in the CLT off the tails.

Our proof employs the expansion of random variables on $[0,1]$ (equipped with Lebesgue measure, on Borel sets) with respect to the Haar basis. The Haar basis is the simplest orthonormal system for $L^2([0,1])$. By considering random variables in $L^2([0,1])$, this proof is consistent with the assumption of finite variance. Therefore, the Haar basis is a natural tool for proving the CLT.

The proof proceeds as follows: Given an i.i.d. sequence of random variables on a probability space, we construct an i.i.d. sequence on $[0,1]$ with the Borel sigma algebra and Lebesgue measure having the same sequence of distributions. As the new sequence of random variables is defined on $[0,1]$ and also has finite variance, we then expand this sequence with respect to the Haar basis.

We then reduce the problem of showing weak convergence of this new sequence of random variables to the case where the Haar expansions are truncated to have only M terms, for some finite M which will be chosen to accomplish certain other objectives. (Lemma 1) These truncated Haar expansions each have $m = 2^{M+1}$ possible outcomes. Next, in Proposition 1, we show that the sum of Haar expansions having only M terms is in fact the projection of a multinomial random variable.

In Lemma 2, we identify (via the constant b_0 introduced there) the tails of the multinomial random variable. After cutting off these tails, we compute the probabilities for the multinomial distribution using Stirlings’s formula and Taylor series approximation (Lemma 3). The appearance of the Gaussian density on the multinomial side can be seen in this step.

On the Gaussian side, we express a standard normal random variable as a sum of m independent normal random variables with coefficients being the outcomes of the truncated Haar expansion. We then apply Fubini’s Theorem to reduce by one dimension the expression for the expected value on the Gaussian side as an integral over a hyperplane in \mathbb{R}^m (Lemma 4). In Lemma 5, we identify (via the constant b_1 introduced there) the tails of the Gaussian. After cutting off these tails, we approximate the integral by a Riemann sum. In Proposition 2, we pull together the results of Lemmas 4 and 5. The Riemann and the multinomial sums match perfectly.

In Theorem 1, by bounding the function f by its sup norm, we estimate the sum of the absolute values of the differences between the multinomial and Gaussian probabilities, establishing strong convergence off the tails. It is here that we also obtain the rate of convergence of $n^{-1/2}$ and

the constant for convergence of $\frac{2m^2}{3\sqrt{2\pi}}$, also off the tails. In both instances, the restriction to “off the tails” arises since our truncations (of the Haar expansions, the multinomial sum, and the Gaussian Riemann sum) are based Chebyshev’s inequality, in which coarseness is the price of its generality. Finally, in Theorem 2, we pull together the preceding results to prove the CLT.

2. PRELIMINARY ESTIMATES

Let $\epsilon > 0$. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded, continuous function. Let Z be a random variable on a probability space (Ω, \mathcal{F}, P) . We may assume that $E(Z)=0$ and $\text{var}(Z)=1$. Define the quantile of Z to be the function $X : [0, 1] \rightarrow \mathbb{R}$ defined by

$$X(x) := \inf\{y \in \mathbb{R} | P(Z \leq y) \geq x\}.$$

Then, X is a random variable on $[0, 1]$ (equipped with Lebesgue measure, on Borel sets) having the same distribution as Z .

For $x \in (0, 1)$, let $\epsilon_i(x)$ be the i th bit in the binary expansion of x (for dyadic rationals, choose the expansion with the tail of 0’s). We create the following matrix of binary digits:

$$\begin{pmatrix} \epsilon_1 & \epsilon_3 & \epsilon_6 & \\ \epsilon_2 & \epsilon_5 & \epsilon_9 & \dots \\ \epsilon_4 & \epsilon_8 & \epsilon_{13} & \\ \vdots & & & \end{pmatrix}$$

For all $x \in (0, 1)$, define $P_i(x)$ to have binary expansion given by the i th column of the matrix. Let $X_i(x) := X(P_i(x))$. Then, (X_i) is an i.i.d. sequence of random variables on $[0, 1]$ having the same distribution as X .

Note that by assumption, $X \in L^2([0, 1])$. The Haar basis is the simplest orthonormal system in $L^2([0, 1])$ and consists of the set $S = \{H_{j,k}(x) | 0 \leq j < \infty, 0 \leq k \leq 2^j - 1\} \cup \{\chi_{[0,1]}\}$ where

$$H_{j,k}(x) := \begin{cases} 2^{\frac{j}{2}} & x \in [\frac{k}{2^j}, \frac{k+\frac{1}{2}}{2^j}) \\ -2^{\frac{j}{2}} & x \in [\frac{k+\frac{1}{2}}{2^j}, \frac{k+1}{2^j}) \\ 0 & \text{otherwise} \end{cases}$$

Since $E(X) = 0$ and $\|X\| < \infty$, it follows that

$$X(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} H_{j,k}(x)$$

where $c_{j,k} = \int_0^1 X(x) H_{j,k}(x) dx$. Then,

$$X(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} 2^{\frac{j}{2}} (-1)^{\epsilon_{j+1}(x)} \chi_{\{k\}}(\lfloor 2^j x \rfloor) = \sum_{j=0}^{\infty} 2^{\frac{j}{2}} c_{j, \lfloor 2^j x \rfloor} (-1)^{\epsilon_{j+1}(x)},$$

where, as usual, $\lfloor x \rfloor$ denotes the greatest integer $\leq x$. For $n \geq 1$, define

$$S_n(x) := \sum_{i=1}^n X_i(x) = \sum_{i=1}^n \sum_{j=0}^{\infty} 2^{\frac{j}{2}} c_{j, \lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))},$$

and for $M \geq 1$, define

$$(1) \quad S_{n,M}(x) := \sum_{i=1}^n X_{i,M}(x) = \sum_{i=1}^n \sum_{j=0}^M 2^{\frac{j}{2}} c_{j, \lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))} \text{ and}$$

$$(2) \quad \sigma_M := \sqrt{\sum_{j=0}^M \sum_{k=0}^{2^j-1} c_{j,k}^2}.$$

Lemma 1. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded, continuous function. Then, there exists a positive integer M_0 such that for all $M \geq M_0$:*

$$\left| \int_0^1 f\left(\frac{S_n(x)}{\sqrt{n}}\right) d\lambda(x) - \int_0^1 f\left(\frac{S_{n,M}(x)}{\sigma_M \sqrt{n}}\right) d\lambda(x) \right| < \epsilon(6\|f\|_\infty + 1).$$

Proof. Note that $E\left(\frac{S_n(x)}{\sqrt{n}}\right) = 0$ and $\text{var}\left(\frac{S_n(x)}{\sqrt{n}}\right) = 1$. Let

$$A := \left\{ \left| \frac{S_n}{\sqrt{n}} \right| > L \right\} \text{ and } B_M := \left\{ \left| \frac{S_{n,M}}{\sigma_M \sqrt{n}} \right| > L \right\}.$$

By Chebyshev's inequality,

$$\lambda(A) \leq \frac{1}{L^2} < \epsilon \text{ and } \lambda(B_M) \leq \frac{1}{L^2} < \epsilon$$

for L large. Since f is uniformly continuous on $[-L, L]$, then there exists a $\delta > 0$ such that $x, y \in [-L, L]$ satisfying $|x - y| < \delta$ implies that $|f(x) - f(y)| < \epsilon$. Now, let

$$C_M := \left\{ \left| \frac{S_n}{\sqrt{n}} - \frac{S_{n,M}}{\sigma_M \sqrt{n}} \right| \geq \delta \right\}.$$

There exists an $M_0 \in \mathbb{N}$ such that for all $M \geq M_0$:

$$\text{var}\left(\frac{S_n}{\sqrt{n}} - \frac{S_{n,M}}{\sigma_M \sqrt{n}}\right) \leq (1 - \sigma_M^2) + 2(1 - \sigma_M)\sqrt{1 - \sigma_M^2} + (1 - \sigma_M)^2 < \epsilon\delta^2 \text{ and so}$$

$$\lambda(C_M) \leq \frac{\epsilon\delta^2}{\delta^2} = \epsilon.$$

Now, let $S := A^c \cap B_M^c \cap C_M^c$. Then, $\lambda(S^c) < 3\epsilon$. Hence,

$$\left| E\left(f\left(\frac{S_n}{\sqrt{n}}\right)\right) - E\left(f\left(\frac{S_{n,M}}{\sigma_M \sqrt{n}}\right)\right) \right| \leq 2\|f\|_\infty \lambda(S^c) + \epsilon \lambda(S) \leq \epsilon(6\|f\|_\infty + 1).$$

□

Let

$$X_M(x) := \frac{1}{\sigma_M} \sum_{j=0}^M 2^{\frac{j}{2}} c_{j, \lfloor 2^j x \rfloor} (-1)^{\epsilon_{j+1}(x)}$$

for $x \in [0, 1]$. Below, we investigate the properties of this random variable. Note that X_M is a random variable which depends on $(\epsilon_1, \dots, \epsilon_{M+1})$. From now on, we will let $m := 2^{M+1}$ for notational convenience. Thus, X_M is constant on dyadic intervals of length $2^{-(M+1)} (= \frac{1}{m})$. Let o_1, \dots, o_m denote the m values. It follows that

$$\sum_{i=1}^m o_i = 0 \text{ and } \sum_{i=1}^m o_i^2 = m$$

as $E(X_M) = 0$ and $\text{var}(X_M) = 1$.

We will now take a closer look at the random variable $S_{n,M}$ of Equation (1). Each $X_{i,M}$ is a random variable with the m possible outcomes o_1, \dots, o_m . Let K_i be the random variable which denotes the number of times the outcome o_i is observed among n independent trials, having outcomes k_i . Then,

$$(3) \quad S_{n,M}(x) = K_1(x)o_1 + \dots + K_m(x)o_m$$

where $K_1 + \dots + K_m = n$.

Note that $S_{n,M}$ is a scalar product of an m -nomial random variable and the vector of outcomes. Since each outcome has probability $\frac{1}{m}$ and the trials are independent,

$$\lambda(\{x \in (0,1) : K_1(x) = k_1, \dots, K_m(x) = k_m\}) = \binom{n}{k_1, \dots, k_m} \left(\frac{1}{m}\right) \text{ and}$$

$$E\left(f\left(\frac{S_{n,M}(x)}{\sqrt{n}}\right)\right) = \sum_{\substack{k_1=0 \\ \vdots \\ k_1+\dots+k_m=n}}^n \dots \sum_{k_m=0}^n \frac{1}{m} \binom{n}{k_1, \dots, k_m} f\left(\frac{\sum_{i=1}^m k_i o_i}{\sqrt{n}}\right).$$

Proposition 1. *Let X be a random variable on $[0,1]$ having mean 0 and variance 1, let*

$$S_n(x) := \sum_{i=1}^n X(P_i(x)),$$

and for each $M > 0$, let $S_{n,M}$ be as in Equation (1) and σ_M be as in Equation (2). Then, for each bounded and continuous $f : \mathbb{R} \rightarrow \mathbb{R}$, and each $\epsilon > 0$, there exists $M_0 \in \mathbb{N}$ such that for all $M \geq M_0$,

$$\left| \int_0^1 f\left(\frac{S_n(x)}{\sqrt{n}}\right) d\lambda(x) - \int_0^1 f\left(\frac{S_{n,M}(x)}{\sigma_M \sqrt{n}}\right) d\lambda(x) \right| < \epsilon(6\|f\|_\infty + 1).$$

Proof. The theorem follows from Lemma 1 and the discussion following Lemma 1. □

The following lemma allows us to cut off the tails from the multinomial random variable. Consequently, we prepare the ground for the usage of Taylor's formula. The tails of the multinomial random variable consist of all $(k_1, \dots, k_m) \in \{0, 1, \dots, n\}^m$ such that $k_1 + \dots + k_m = n$ and $k_i \notin [\lfloor \frac{n}{m} \rfloor - \lfloor b\sqrt{n} \rfloor, \lfloor \frac{n}{m} \rfloor + \lfloor b\sqrt{n} \rfloor]$ for all $1 \leq i \leq m-1$.

Lemma 2. *Let*

$$q(n, k_1, \dots, k_m) := \left(\frac{1}{m}\right)^n \binom{n}{k_1, \dots, k_m} f\left(\frac{\sum_{i=1}^m k_i o_i}{\sqrt{n}}\right).$$

Then, there exists a b_0 such that for all $b \geq b_0$:

$$\left| \sum_{\substack{k_1=1 \\ \vdots \\ k_1+\dots+k_m=n}}^n \dots \sum_{k_m=0}^n q(n, k_1, \dots, k_m) - \sum_{\substack{k_1=\lfloor \frac{n}{m} \rfloor - \lfloor b\sqrt{n} \rfloor \\ \vdots \\ k_1+\dots+k_m=n}}^{\lfloor \frac{n}{m} \rfloor + \lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{k_{m-1}=\lfloor \frac{n}{m} \rfloor - \lfloor b\sqrt{n} \rfloor \\ \vdots \\ k_1+\dots+k_m=n}}^{\lfloor \frac{n}{m} \rfloor + \lfloor b\sqrt{n} \rfloor} q(n, k_1, \dots, k_m) \right| < \epsilon \|f\|_\infty.$$

Proof. Recall that K_i is the random variable which denotes the number of times the outcome o_i is observed, having values k_i . As each K_i is a binomial random variable, we have $E(K_i) = \frac{n}{m}$ and $\text{var}(K_i) = n(\frac{1}{m})(1 - \frac{1}{m})$. By Chebyshev's inequality,

$$\lambda\left(|K_i - \frac{n}{m}| \geq b\sqrt{n}\right) \leq \frac{(1 - \frac{1}{m})}{b^2 m} \leq \frac{1}{b^2 m}.$$

Then, there exists a b_0 such that for all $b \geq b_0$:

$$\lambda\left(|K_i - \frac{n}{m}| \geq b\sqrt{n} \text{ for some } 1 \leq i \leq m\right) \leq \frac{1}{b^2} < \epsilon.$$

□

In the following lemma, we will use Stirling's formula and Taylor series to approximate the probabilities for the multinomial distribution. For use here and in the proof of Theorem 2, we define some functions. For $n > 0$, we let:

$$(4) \quad d_n := \sqrt{m} \left(\frac{m}{2n\pi}\right)^{\frac{m-1}{2}}.$$

For $n > 0$ and integers, j_1, \dots, j_m whose sum is 0, we let:

$$(5) \quad H(n, j_1, \dots, j_m) := \left(\frac{m^2}{4n^2} - \frac{m}{2n} \right) \sum_{i=1}^m j_i^2 + \left(\frac{m^2}{6n^2} - \frac{m^3}{6n^3} \right) \sum_{i=1}^m j_i^3 - \frac{m^3}{3n^3} \sum_{i=1}^m j_i^4 + O(n^{-1}), \text{ and}$$

$$(6) \quad p(n, j_1, \dots, j_m) := d_n e^{H(n, j_1, \dots, j_m)}.$$

Lemma 3. Let $j_i = k_i - \lfloor \frac{n}{m} \rfloor$, and suppose that $- \lfloor b\sqrt{n} \rfloor \leq j_i \leq \lfloor b\sqrt{n} \rfloor$ for $1 \leq i \leq m-1$ and $j_m = -j_1 - \dots - j_{m-1}$ and $n \geq b^2 m^2$. Then,

$$\frac{1}{m^n} \binom{n}{k_1, \dots, k_m} = \left(1 + O\left(\frac{1}{n}\right) \right) p(n, j_1, \dots, j_m).$$

Proof. Set

$$l(n, k_1, \dots, k_m) := \frac{1}{m^n} \binom{n}{k_1, \dots, k_m}.$$

By Stirling's Formula, we have

$$l(n, k_1, \dots, k_m) = \frac{(1 + O(\frac{1}{n})) (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}}}{(2\pi)^{\frac{m}{2}} m^n k_1^{(k_1+\frac{1}{2})} \dots k_m^{(k_m+\frac{1}{2})}}.$$

Letting $k_i = \frac{n}{m} + j_i$ for $1 \leq i \leq m$,

$$\begin{aligned} l(n, k_1, \dots, k_m) &= \frac{(1 + O(\frac{1}{n})) (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}}}{(2\pi)^{\frac{m}{2}} m^n (\frac{n}{m} + j_1)^{(\frac{n}{m}+j_1+\frac{1}{2})} \dots (\frac{n}{m} + j_m)^{(\frac{n}{m}+j_m+\frac{1}{2})}} \\ &= \frac{(1 + O(\frac{1}{n})) m^{\frac{m}{2}}}{(2\pi)^{\frac{m-1}{2}} n^{\frac{m-1}{2}} (1 + \frac{mj_1}{n})^{(\frac{n}{m}+j_1+\frac{1}{2})} \dots (1 + \frac{mj_m}{n})^{(\frac{n}{m}+j_m+\frac{1}{2})}}. \end{aligned}$$

For all $1 \leq i \leq m$, we set

$$a(n, m, i) := (1 + \frac{mj_i}{n})^{\frac{n}{m}+j_i+\frac{1}{2}} = e^{(\frac{n}{m}+j_i+\frac{1}{2}) \ln(1+\frac{mj_i}{n})}.$$

Using a Taylor series approximation, we have the following for n large enough (as j_i is bounded by $O(\sqrt{n})$):

$$\begin{aligned} a(n, m, i) &= \exp \left(\left(\frac{n}{m} + j_i + \frac{1}{2} \right) \left(\frac{mj_i}{n} - \frac{m^2 j_i^2}{2n^2} + \frac{m^3 j_i^3}{3n^3} + O(n^{-1}) \right) \right) \\ &= \exp \left(j_i + \frac{mj_i^2}{2n} + \frac{mj_i}{2n} - \frac{m^2 j_i^3}{6n^2} - \frac{m^2 j_i^2}{4n^2} + \frac{m^3 j_i^4}{3n^3} + \frac{m^3 j_i^3}{6n^3} + O(n^{-1}) \right). \end{aligned}$$

Therefore, we have

$$(1 + \frac{mj_1}{n})^{\frac{n}{m}+j_1+\frac{1}{2}} \dots (1 + \frac{mj_m}{n})^{\frac{n}{m}+j_m+\frac{1}{2}} = e^{H(n, j_1, \dots, j_m)}, \text{ and so}$$

$$l(n, j_1, \dots, j_m) = \left(1 + O\left(\frac{1}{n}\right) \right) p(n, j_1, \dots, j_m), \text{ as required.}$$

□

Now we consider the Gaussian side. Let Y_1, \dots, Y_n be i.i.d. standard normal random variables. Then, by the properties of i.i.d. normal random variables and by Lemma 2,

$$\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i Y_i$$

is a standard normal random variable.

Lemma 4. Let S be the hyperplane in \mathbb{R}^m given by $\sum_{i=1}^m y_i = 0$. Then,

$$E \left(f \left(\sum_{i=1}^m \frac{o_i Y_i}{\sqrt{m}} \right) \right) = E_{\mathcal{L}(Z)} \left(f \left(\sum_{i=1}^m \frac{o_i Y_i}{\sqrt{m}} \right) \right),$$

where $\mathcal{L}(Z)$ is a standard Gaussian $m-1$ dimensional measure on the hyperplane S . Moreover,

$$E_{\mathcal{L}(Z)} \left(f \left(\frac{\sum_{i=1}^m o_i Y_i}{\sqrt{m}} \right) \right) = \frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \int_S \dots \int f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) e^{(-\frac{1}{2} \sum_{i=1}^m y_i^2)} dy_1 \dots dy_{m-1}.$$

Proof. Set

$$o = \begin{bmatrix} o_1 \\ \vdots \\ o_m \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Then, the vector projection of Y in direction of the vector u is given by

$$V := \left(\frac{1}{m} \sum_{i=1}^m Y_i \right) u = \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m Y_i \right) \frac{u}{\sqrt{m}}.$$

Thus, V is a one dimensional standard normal on the line through the origin and orthogonal to the hyperplane S . Viewing V and $Z := Y - V$ as vector valued random variables, V and Z are independent as random variables in addition to being orthogonal as vectors. This can be verified by checking that all components of Z are independent of all components of V . Since all components of V are equal to $\bar{Y} := \frac{1}{m} \sum_{i=1}^m Y_i$, and since the components of $Y - V$ are $Y_i - \bar{Y}$, then the independence of the Gaussian random variables follows as

$$E((Y_i - \bar{Y})\bar{Y}) = E(Y_i \bar{Y}) - E(\bar{Y})^2 = \frac{1}{m} - \frac{1}{m} = 0$$

for all $i = 1, \dots, m$. Note that $o^T Y = o^T (Y - V) + o^T V$. Since $o^T u = 0$,

$$f \left(\frac{\sum_{i=1}^m o_i Y_i}{\sqrt{m}} \right) = f \left(\frac{1}{\sqrt{m}} o^T Y \right) = f \left(\frac{1}{\sqrt{m}} o^T Z \right).$$

That is, $f \left(\frac{1}{\sqrt{m}} o^T Z \right)$ does not depend on V . Note that since the components of Z satisfy $\sum_{i=1}^m (Y_i - \bar{Y}) = 0$, $\mathcal{L}(Z)$, the law of Z , is a standard Gaussian $m-1$ dimensional measure on the hyperplane S . From the independence of Z and V , and Fubini's Theorem, it follows that

$$Ef \left(\frac{1}{\sqrt{m}} o^T Y \right) = E_{\mathcal{L}(Z)} E_{\mathcal{L}(V)} f \left(\frac{1}{\sqrt{m}} o^T Z \right) = E_{\mathcal{L}(Z)} f \left(\frac{1}{\sqrt{m}} o^T Z \right).$$

Since the density of Y is given by

$$\frac{1}{(\sqrt{2\pi})^m} \exp \left(-\frac{1}{2} y^T y \right) = \frac{1}{(\sqrt{2\pi})^{m-1}} \exp \left(-\frac{1}{2} z^T z \right) \frac{1}{(\sqrt{2\pi})} \exp \left(-\frac{1}{2} (\bar{y})^2 \right),$$

where y , z , and \bar{y} are the realizations of Y , Z , and \bar{Y} , respectively, it follows that

$$E_{\mathcal{L}(Z)} f \left(\frac{1}{\sqrt{m}} o^T Z \right) = \frac{1}{(\sqrt{2\pi})^{m-1}} \int_S f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) \exp \left(-\frac{1}{2} \sum_{i=1}^m y_i^2 \right) dS.$$

That is, the expected value with respect to $\mathcal{L}(Z)$ is a surface integral over the hyperplane S . In arriving at the last equality we have used that $\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i = 0$ on S . By projecting S onto the $y_m = 0$ plane, we have

$$E_{\mathcal{L}(Z)} f \left(\frac{1}{\sqrt{m}} o^T Z \right) = \frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \int \dots \int f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) e^{(-\frac{1}{2} \sum_{i=1}^m y_i^2)} dy_1 \dots dy_{m-1}$$

where $y_m = -\sum_{i=1}^{m-1} y_i$. □

Lemma 5. *Let*

$$I := \frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \int \dots \int_{y_1+y_2+\dots+y_m=0} f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) e^{(-\frac{1}{2} \sum_{i=1}^m y_i^2)} dy_1 \dots dy_{m-1}.$$

Then, there exist $n_0 \in \mathbb{N}$ and $b_1 > 0$ such that for all $n \geq n_0$ and for all $b \geq b_1$,

$$\left| I - \frac{m^{\frac{m}{2}}}{(\sqrt{2\pi})^{m-1} n^{\frac{m-1}{2}}} \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+j_2+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} f \left(\frac{1}{\sqrt{n}} \sum_{i=1}^m o_i j_i \right) e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \right| < \epsilon(\|f\|_\infty + 1).$$

Proof. Let

$$C := \{y \in \mathbb{R}^m : y_i \in [-b\sqrt{m}, b\sqrt{m}] \text{ for all } 1 \leq i \leq m-1 \text{ and } y_m = -(y_1 + \dots + y_{m-1})\}.$$

We have

$$\frac{1}{(\sqrt{2\pi})^{m-1}} \int_S \exp \left(-\frac{1}{2} \sum_{i=1}^m y_i^2 \right) dS = \frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \int \dots \int_{y_1+y_2+\dots+y_m=0} \exp \left(-\frac{1}{2} \sum_{i=1}^m y_i^2 \right) dy_1 \dots dy_{m-1} = 1.$$

Thus, there exists a b_1 such that for all $b \geq b_1$,

$$\frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \left(\int \dots \int_{y \in C^c} e^{-\frac{1}{2} \sum_{i=1}^m y_i^2} dy_1 \dots dy_{m-1} \right) < \epsilon, \text{ and so:}$$

$$\left| I - \frac{\sqrt{m}}{(\sqrt{2\pi})^{m-1}} \int \dots \int_{y \in C} f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) e^{(-\frac{1}{2} \sum_{i=1}^m y_i^2)} dy_1 \dots dy_{m-1} \right| < \epsilon \|f\|_\infty.$$

Suppose that $b \geq b_1$. Let

$$I_b := \frac{\sqrt{m}}{(2\pi n)^{\frac{m-1}{2}}} \int \dots \int_{y \in C} f \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i y_i \right) e^{(-\frac{1}{2} \sum_{i=1}^m y_i^2)} dy_1 \dots dy_{m-1}.$$

Then,

$$I_b = \lim_{n \rightarrow \infty} \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+j_2+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} f \left(\frac{1}{\sqrt{n}} \sum_{i=1}^m o_i j_i \right) e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)}.$$

Hence, there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\left| I - \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+j_2+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} f \left(\frac{1}{\sqrt{n}} \sum_{i=1}^m o_i j_i \right) e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \right| < \epsilon(\|f\|_\infty + 1).$$

□

Proposition 2. *Let Y be a standard normal random variable. Then, for each $\epsilon > 0$, for each bounded, continuous $f : \mathbb{R} \rightarrow \mathbb{R}$, there exists $n_0 \in \mathbb{N}$, $b_1 > 0$ such that for all $n \geq n_0$ and for all $b \geq b_1$,*

$$\left| E(f(Y)) - \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + j_2 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + j_2 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} f\left(\frac{1}{\sqrt{n}} \sum_{i=1}^m o_i j_i\right) e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \right| < \epsilon(\|f\|_\infty + 1).$$

Proof. The statement is immediate from Lemma 4 and Lemma 5. □

3. MAIN RESULTS

Theorem 2 gives our proof of the CLT. The proof appeals to Theorem 1, which also establishes strong convergence off the tails and gives a rate and constant of convergence there; the rate is $n^{-1/2}$ and the constant of convergence is $\frac{2m^2}{3\sqrt{2\pi}}$. These only hold off the tails as we have truncated the Haar expansions, the multinomial sum, and the Gaussian Riemann sum. Theorem 1 combines the results of Propositions 1 and 2. From now on, let $b \geq \max\{b_0, b_1\}$, where the former is as in Lemma 2 and the latter is as in the proof of Lemma 5.

Theorem 1. *Let*

$$D_n := \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \left| \left(\frac{1}{m^n} \binom{n}{\frac{n}{m} + j_1, \dots, \frac{n}{m} + j_m} - \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \right) \right|.$$

Then,

$$D_n \leq \frac{2m^2}{3\sqrt{2\pi n}} + O(n^{-1}).$$

Proof. By Lemma 3,

$$D_n = \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \left| e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \left(e^{G(n, j_1, \dots, j_m) + O(n^{-1})} - 1 \right) \right|$$

where $G(n, j_1, \dots, j_m) = \left(\frac{m^2}{4n^2}\right) \sum_{i=1}^m j_i^2 + \left(\frac{m^2}{6n^2} - \frac{m^3}{6n^3}\right) \sum_{i=1}^m j_i^3 - \frac{m^3}{3n^3} \sum_{i=1}^m j_i^4$. Then,

$$\begin{aligned} D_n &= \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} |e^{G(n, j_1, \dots, j_m) + O(n^{-1})} - 1| \leq \\ &\quad \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} |G(n, j_1, \dots, j_m) + O(n^{-1})|. \end{aligned}$$

All of the terms which decay more quickly than $\frac{1}{\sqrt{n}}$ are absorbed into the error term $O(\frac{1}{n})$. It remains to compute the constant $C(m)$. Define

$$L_n := \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \left(\frac{m^2}{6n^2} \right) \sum_{\substack{j_{m-1} = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1 = -\lfloor b\sqrt{n} \rfloor \\ j_1 + \dots + j_m = 0}}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \left(\sum_{i=1}^{m-1} |j_i|^3 + |j_m|^3 \right).$$

We let

$$E_n := \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \left(\frac{m^2}{6n^2} \right) \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} \left(\sum_{i=1}^{m-1} |j_i|^3 \right) \leq$$

$$\frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \left(\frac{m^2}{6n^2} \right) \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^{m-1} \frac{j_i^2}{n}\right)} \left(\sum_{i=1}^{m-1} |j_i|^3 \right).$$

Approximating the sum by an integral, we have

$$E_n < (m-1)\sqrt{m} \cdot \frac{m^2}{6n^2} \cdot \left(\frac{n}{m} \right)^{3/2} E(|X|^3) + O(n^{-1})$$

where X is a standard normal random variable. Thus,

$$E_n < \frac{2}{\sqrt{2\pi}} \frac{m(m-1)}{3\sqrt{n}} + O(n^{-1}).$$

Now, consider

$$F_n := \left(\frac{m^2}{6n^2} \right) \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \sum_{j_{m-1}=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n}\right)} |j_m|^3.$$

By maximizing $e^{-\frac{1}{2}x^2}|x|^3$,

$$F_n \leq \left(\frac{m^2}{6n^2} \right) \frac{m^{\frac{m}{2}}}{(2\pi n)^{\frac{m-1}{2}}} \left(\frac{n}{m} \right)^{3/2} \sum_{j_{m-1}=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{j_1=-\lfloor b\sqrt{n} \rfloor}^{\lfloor b\sqrt{n} \rfloor} e^{-\left(\frac{m}{2} \sum_{i=1}^{m-1} \frac{j_i^2}{n}\right)} e^{-3/2} 3^{3/2}.$$

Approximating the sum by an integral, we have

$$F_n < \sqrt{m} \cdot \left(\frac{m^2}{6n^2} \right) \cdot \left(\frac{n}{m} \right)^{3/2} e^{-3/2} 3^{3/2} + O(n^{-1}).$$

Thus,

$$F_n < \frac{m}{6\sqrt{n}} e^{-3/2} 3^{3/2} + O(n^{-1}).$$

Hence, we have

$$E_n + F_n < \frac{2}{\sqrt{2\pi}} \cdot \frac{m(m-1)}{3\sqrt{n}} + \frac{m}{6\sqrt{n}} e^{-3/2} 3^{3/2} + O(n^{-1}) < \frac{2m^2}{3\sqrt{2\pi n}} + O(n^{-1}).$$

It then follows that

$$D_n \leq \frac{2m^2}{3\sqrt{2\pi n}} + O(n^{-1}).$$

□

Finally, we prove the CLT using Lemmas 1 - 5 and Theorem 1. Recall from Lemma 2 that:

$$q(n, k_1, \dots, k_m) := \left(\frac{1}{m} \right)^n \binom{n}{k_1, \dots, k_m} f\left(\frac{\sum_{i=1}^m k_i o_i}{\sqrt{n}} \right).$$

Recall from Lemma 3 that:

$$d_n := \sqrt{m} \left(\frac{m}{2n\pi} \right)^{\frac{m-1}{2}}.$$

For $n > 0$ and integers, j_1, \dots, j_m whose sum is 0:

$$H(n, j_1, \dots, j_m) := \left(\frac{m^2}{4n^2} - \frac{m}{2n} \right) \sum_{i=1}^m j_i^2 + \left(\frac{m^2}{6n^2} - \frac{m^3}{6n^3} \right) \sum_{i=1}^m j_i^3 - \frac{m^3}{3n^3} \sum_{i=1}^m j_i^4 + O(n^{-1}), \text{ and}$$

$$p(n, j_1, \dots, j_m) := d_n e^{H(n, j_1, \dots, j_m)}.$$

Theorem 2. Let (X_i) be a sequence of i.i.d. random variables with mean μ and variance σ^2 . Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded, continuous function. Then, for each $\epsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that

$$\left| E \left(f \left(\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \right) \right) - E(f(Y)) \right| < \epsilon(9\|f\|_\infty + 2)$$

for all $n \geq n_1$, where Y is a standard normal random variable.

Proof. By Lemma 1, we reduce the problem to dealing with the projection of a multinomial random variable and we have

$$\Delta_n := \left| E \left(f \left(\frac{S_n}{\sqrt{n}} \right) \right) - E(f(Y)) \right| < \left| E \left(f \left(\frac{S_{n,M}}{\sigma_M \sqrt{n}} \right) \right) - E(f(Y)) \right| + \epsilon(6\|f\|_\infty + 1) =$$

$$\left| \sum_{\substack{k_1=0 \\ k_1+\dots+k_m=n}}^n \dots \sum_{\substack{k_m=0 \\ k_1+\dots+k_m=n}}^n q(n, k_1, \dots, k_m) - E(f(Y)) \right| + \epsilon(6\|f\|_\infty + 1).$$

By Lemma 2, we cut off the tails of the multinomial random variable to obtain

$$\Delta_n < \left| \sum_{\substack{k_1=\lfloor \frac{n}{m} \rfloor - \lfloor b\sqrt{n} \rfloor \\ k_1+\dots+k_m=n}}^{\lfloor \frac{n}{m} \rfloor + \lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{k_{m-1}=\lfloor \frac{n}{m} \rfloor - \lfloor b\sqrt{n} \rfloor \\ k_1+\dots+k_m=n}}^{\lfloor \frac{n}{m} \rfloor + \lfloor b\sqrt{n} \rfloor} q(n, k_1, \dots, k_m) - E(f(Y)) \right| + \epsilon(7\|f\|_\infty + 1).$$

By Lemma 3, we further simplify the multinomial sum to obtain

$$\Delta_n < \left| \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} p(n, j_1, \dots, j_m) f \left(\frac{\sum_{i=1}^m j_i o_i}{\sqrt{n}} \right) - E(f(Y)) \right| + \epsilon(7\|f\|_\infty + 1).$$

Writing Y as a sum of m independent standard normal random variables, it follows that $\frac{1}{\sqrt{m}} \sum_{i=1}^m o_i Y_i \stackrel{d}{=} N(0, 1)$. By Lemma 4,

$$E \left(f \left(\sum_{i=1}^m \frac{o_i Y_i}{\sqrt{m}} \right) \right) = E_{\mathcal{L}(Z)} \left(f \left(\sum_{i=1}^m \frac{o_i Y_i}{\sqrt{m}} \right) \right).$$

By Lemma 5, we approximate the Gaussian integral by a Riemann sum. This approximation allows us to match the multinomial side and the Gaussian side and apply Theorem 1:

$$\Delta_n < d_n \left| \sum_{\substack{j_{m-1}=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} \dots \sum_{\substack{j_1=-\lfloor b\sqrt{n} \rfloor \\ j_1+\dots+j_m=0}}^{\lfloor b\sqrt{n} \rfloor} f \left(\frac{\sum_{i=1}^m j_i o_i}{\sqrt{n}} \right) \left(e^{H(n, j_1, \dots, j_m)} - e^{-\left(\frac{m}{2} \sum_{i=1}^m \frac{j_i^2}{n} \right)} \right) \right| + \epsilon(8\|f\|_\infty + 2) <$$

$$\epsilon(8\|f\|_\infty + 2) + \frac{2m^2\|f\|_\infty}{3\sqrt{2\pi n}} + O(n^{-1}) < \epsilon(9\|f\|_\infty + 2)$$

for all $n \geq n_1$. □

REFERENCES

- [1] BERRY, ANDREWS C. (1941) "The Accuracy of the Gaussian Approximation to the Sum of Independent Variates." Transactions of the American Mathematical Society 49 (1): 122-136.
- [2] DALANG, R.C. (2005) "Une démonstration élémentaire du théorème central limite." Elemente der Mathematik, vol. 60, no. 1, pp. 1-9.
- [3] ESSEEN, CARL-GUSTAV (1942) "On the Liapunoff limit of error in the theoretical probability." Arkiv for matematik, astronomi och fysik A28:1-19.
- [4] ESEEN, CARL-GUSTAV (1956) "A moment inequality with an application to the central limit theorem." Skand. Aktuarietidskr. 39: 160-170.

- [5] FELLER, W. (1935). "Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung." *Mathematische Zeitschrift* **40**, 521-559.
- [6] FISCHER, H. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer, 2011.
- [7] GARMIRIAN, P. The Central Limit Theorem and the Estimation for the Concentration of Measure for Fractional Brownian Motion. Dissertation, Lehigh University (2013).
- [8] LÉVY, P. (1935). "Propriétés asymptotiques des sommes des ariables aleatoires enchainées." *Bulletin de la Societé mathématique de France* (2) 59, 1-32.
- [9] LÉVY, P. (1935). "Propriétés asymptotiques des sommes des variables aléatoires independantes ou enchainées." *Journal de mathématiques pures et appliquées* 14, 347-402.
- [10] TROTTER, H.F. (1959) "An Elementary Proof of the Central Limit Theorem." *Arch. Math.* 10: 226-234.
- [11] ZONG, ZHAOJUN AND FENG HU (2013) "A New Proof of the Central Limit Theorem for i.i.d. random variables." *Abstract and Applied Analysis*, vol. 2013, Article ID 294910, 5 pages.

DEPARTMENT OF MATHEMATICS, LEHIGH UNIVERSITY, 14 EAST PACKER AVENUE, BETHLEHEM, PA 18015
E-mail address: `vd00@lehigh.edu`

DEPARTMENT OF MATHEMATICS, TUFTS UNIVERSITY, BOSTON AVENUE, MEDFORD, MA 02155
E-mail address: `Patricia.Garmirian@tufts.edu` (corresponding author)